A Rapid Genotyping Assay for Segregating Human Olfactory Receptor Pseudogenes

Craig S. Hinkley and Lindita Ismaili*

Department of Biological Sciences, Kingsborough Community College, Brooklyn, New York 11235, USA

Variation in odor perception between individuals is initiated by binding of "odorant" molecules to olfactory receptors (ORs) located in the nasal cavity. To determine the mechanism for variation in odor perception, identification of specific ligands for a large number of ORs is required. However, it has been difficult to identify specific ligands, and ligands have been identified for only 2–3% of the hundreds of mammalian ORs. One way to increase the number of identified ligands is to take advantage of >60 human OR genes that are segregating as a result of a single nucleotide polymorphism, between a functional intact allele and a nonfunctional pseudogene allele. Potential ligands for these ORs can be identified by correlating odor perception of an individual with their genotype [intact/intact (I/I) vs. pseudogene/pseudogene (P/P)] for an OR gene. For this type of study, genotypes must be determined for a large number of individuals. We have developed a PCR-based assay to distinguish between the intact and pseudogene alleles of 49 segregating human OR genes and to determine an individual's genotype for these genes. To facilitate rapid determination of genotypes for a large number of individuals, the assay uses a small number of simple steps and equipment commonly found in most molecular biology and biochemistry laboratories. Although this assay was developed to distinguish between polymorphisms in OR genes, it can easily be adapted for use in distinguishing single nucleotide polymorphisms in any gene or chromosomal locus.

KEY WORDS: cheek cell DNA, ligands, PCR, single nucleotide polymorphism

INTRODUCTION

Initiation of odor perception occurs through binding of odorant molecules to ORs located in olfactory sensory neurons of the nasal epithelium. Each olfactory sensory neuron expresses only one allele of a single OR gene; therefore, odorant binding is determined by the properties of individual ORs.[1] These receptors are members of a large superfamily of seven-transmembrane domain GPCRs that have been conserved throughout vertebrate evolution.[2–5] Studies with the rat OR-17 and other ORs have shown that each OR can recognize several structurally related ligands, each of which can be recognized by multiple OR types.[6,7] In this scheme, each ligand has a unique code combination of receptors responding to it, which the brain interprets as different odors.[8,9] Methods to identify ligands have included studies of endogenous ORs, cloned ORs expressed in cultured cells, surface expression of ORs in heterologous cells, and odor perception tests.[6,8–11] Although these studies have yielded valuable insights into OR function, identifying ligands for ORs has been surprisingly difficult, and ligands have been identified for only 2–3% of the hundreds of mammalian ORs.

Humans have 800–900 OR genes, of which ~60% are nonfunctional pseudogenes.[4,12] However, >60 of these pseudogenes are still segregating between a functional intact allele and a nonfunctional pseudogene allele in the human population. Most of the segregating pseudogenes are nonfunctional, as they contain a deleterious single nucleotide substitution or deletion in their coding regions.[13] Menashe et al.[14] suggested that ligands for these segregating human OR pseudogenes possibly can be identified by comparing the genotypes of individuals—I/I, I/P, and P/P—with their phenotypes; the detection of different odorants during odorant perception tests. They then used a genotype/phenotype comparison approach to identify isovaleric acid as a ligand for human OR11H7. For a genotype/phenotype study, it is important to identify large numbers of individuals with each genotype. In one-half of the 52 pseudogenes used in the study by Menashe et al.[14], the frequency of one allele was <25%, making it likely that a large number of individuals will need to be genotyped for

*ADDRESS CORRESPONDENCE TO: Craig S. Hinkley, Department of Biological Sciences, Kingsborough Community College, 2001 Oriental Blvd., Brooklyn, NY 11235, USA (Phone: 718-368-5783; Fax: 718-368-4873; E-mail: chinkley@kbcc.cuny.edu).

The authors declare no conflicts of interest. Neither funding source was involved with the research or paper.
doi: 10.7171/jbt.12-2303-001

these pseudogenes to obtain a large enough sample size of individuals with each genotype. In this paper, a PCR-based assay is presented that can be used to identify individuals quickly with I/I, I/P, and P/P genotypes for segregating human OR pseudogenes.

MATERIALS AND METHODS

Primer Design

The Human Olfactory Data Explorer (HORDE) database was used to identify the human OR genes that are still segregating between an intact allele and a pseudogene allele.[15] The intact gene and pseudogene polymorphisms were identified from the work of Menashe et al.[13,14] and the HORDE database.[15] For each segregating human OR gene, a primer set was designed that could amplify the intact and pseudogene allele, as well as produce an amplified DNA fragment containing the corresponding allele polymorphism. Determination of an individual's genotype is possible, as each primer set was designed so that a unique restriction site is present in the amplified DNA fragment from only one of the two alleles. Therefore, digestion of the amplified DNA with the unique restriction enzyme will result in cutting only the DNA amplified from the intact gene or the DNA amplified from the pseudogene, depending on the presence of the unique restriction site.

In the best-case scenario, the intact gene or pseudogene polymorphism formed a unique restriction site. This was determined using dCAPs Finder 2.0 software and allowing zero mismatches in the primer.[16] When no restriction site was formed using solely the existing polymorphism, a primer was designed that would generate a unique restriction site in the DNA amplified from the intact gene or pseudogene. This primer was designed using dCAPs Finder 2.0 and sequentially increasing the number of mismatches allowed in the primer, to a maximum of five differences, until a unique restriction site was identified. NEBcutter V2.0 software was used to verify a restriction site is unique to the amplified DNA fragment of only one allele of an OR gene.[17] Primer3 software was used to generate the final primer sets.[18] Positions of polymorphisms in each human OR gene and information about digestion of PCR products are indicated in Table 1. Table 2 contains a list of primer sets for PCR amplification of each segregating human OR gene.

Isolation of Genomic DNA

The use of human cheek cells was approved by the Institutional Review Board of Kingsborough Community College (Brooklyn, NY, USA) in accordance with the ethical standards of the U.S. Department of Health and Human Services. Cheek cells were obtained on an anonymous basis, thus no individuals could be identified. All materials obtained from these individuals were destroyed at the end of study. No information that could identify individuals was collected.

This protocol was modified from Bloom.[19] Human cheek cell suspensions were collected from volunteers by mouthwash with 10 ml 0.9% NaCl for 30 s. The cell suspension (1 ml) was transferred to a 1.5-ml Eppendorf tube, and the cells were pelleted by centrifugation for 1 min at 14,000 rpm in a microcentrifuge. The supernatant was decanted until 30–100 μl supernatant remained. Cells were resuspended by pipetting, using the remaining supernatant, and 30 μl of the cell suspension was transferred to a thin-walled (PCR) tube containing 100 μl 10% Chelex 100 resin (200–400 mesh; Bio-Rad Life Sciences, Hercules, CA, USA; Catalogue #142-1253), pH 11. The cells were lysed by heating them for 10 min at 99°C in a thermal cycler, and then the tube containing the lysed cells was shaken well to mix cell contents with the Chelex resin. Note: The use of Chelex resin is the most critical step, as it removes cations that will inhibit PCR amplification. The tubes were then centrifuged for 1 min at 14,000 rpm in a microcentrifuge. The supernatant, containing the genomic DNA, was carefully transferred to a clean, 1.5-ml Eppendorf tube without disturbing the pelleted cell debris and Chelex resin. This genomic DNA may be amplified immediately or stored at −20°C.

PCR Amplification and Detection of Amplified DNA

PCR amplification of the human OR3A1 was performed in 25 μl reactions using PuRe Taq Ready-To-Go PCR beads (GE Healthcare, Waukesha, WI, USA; Catalogue #27-9559-01). The PCR bead was first dissolved in 22.5 μl primer/loading dye mix (6.75 pmol each primer, 34% sucrose, 0.02% cresol red dye), to which 2.5 μl cheek cell DNA was added. Cycling conditions were 94°C for 5 min, 94°C for 30 s, 68°C for 30 s, 72°C for 30 s for 30 cycles, and 72°C for 10 min. The same cycling conditions can be used for each OR gene by changing the annealing temperature for each primer set. After PCR amplification, 10 μl of the amplified DNA was removed and used as an undigested DNA control. Two units of *HinfI* (New England Biolabs, Ipswich, MA, USA; Catalogue #R0155S) were added to the remaining amplified DNA, which was then incubated at 37°C for 30 min. DNA was separated on a 2% agarose gel and visualized by ethidium bromide staining.

RESULTS AND DISCUSSION

The human OR3A1 gene was used to demonstrate the feasibility of the genotyping assay for the segregating human OR pseudogenes. OR3A1 is located on chromosome 17 with an intact allele frequency of 55% and a pseudogene allele frequency of 45%.[13] The intact allele has a "G" at

TABLE 1

Polymorphism and PCR Product Information for Human OR Genes

Gene	DNA polymorphism intact > pseudo[a]	nt[a]	Protein change intact > pseudo[a]	AA[a]	Reference SNP ID[b]	PCR size (bp)[c]	Enzyme[c]	Allele cut[c]	Digestion sizes (bp)[c]
OR1A1	C > T	853	P > S	285	rs769427	135	*Tsp509I*	pseudo	57 + 78
OR1B1	C > T	574	R > stop	192	rs1476860	163	*TaqI*	intact	24 + 139
OR1E3	C > Del(1)	54	P > frameshift	9	rs11377766	154	*SfaNI*	intact	30 + 124
OR1F1	G > A	365	R > H	122	rs61731440	232	*AciI*	intact	63 + 169
OR1P1	A > T	553	K > stop	185	rs7222006	180	*TaqI*	intact	23 + 157
OR1S1	G > A	404	R > H	135	rs1966834	233	*DraIII*	pseudo	113 + 120
OR2A9	C > T	295	Q > stop	99	none	151	*TaqI*	intact	24 + 127
OR2AG1	A > G	125	N > S	42	rs11826041	223	*BsrDI*	intact	69 + 154
OR2J1	C > T	580	Q > stop	194	rs2394517	254	*StyI*	pseudo	106 + 148
OR2L8	A > G	650	Y > C	217	rs4925583	165	*NcoI*	intact	22 + 143
OR2S2	G > A	368	R > H	123	rs2233563	195	*BclI*	pseudo	45 + 150
OR2T11	C > T	355	R > C	119	rs1892443	191	*AciI*	intact	47 + 144
OR3A1	G > A	374	R > Q	125	rs703903	164	*HinfI*	intact	45 + 119
OR4A8	C > T	406	R > stop	136	none	216	*HinfI*	intact	61 + 155
OR4C16	C > T	49	Q > stop	17	rs1459101	200	*HinfI*	intact	22 + 178
OR4E2	A > G	352	M > V	118	rs2874103	150	*BspHI*	intact	18 + 132
OR4K3	C > Del(1)	622	A > frameshift	208	rs5807006	213	*MwoI*	pseudo	85 + 128
OR4X1	T > A	819	Y > stop	273	rs10838851	241	*DraI*	pseudo	85 + 156
OR4X2	C > G	81	Y > stop	27	rs7120775	195	*BsrGI*	intact	38 + 157
OR5AL1	CT > Del(2)	468	L > frameshift	156	rs10633383	159	*HpyCH4III*	pseudo	42 + 117
OR5AR1	C > T	55	Q > stop	19	rs11228710	162	*AvrII*	pseudo	46 + 116
OR5D13	A > G	185	Y > C	62	rs297118	176	*RsaI*	intact	80 + 96
OR5G3	C > Del(1)	372	L > frameshift	124	rs61392974	210	*BstNI*	intact	67 + 143
OR5H6	C > G	433	P > A	145	rs9289564	157	*HindIII*	pseudo	49 + 108
OR5L1	C > T	859	P > S	287	rs12790505	152	*HpaII*	intact	22 + 130
OR5R1	C > T	364	R > C	122	rs6591324	227	*PvuI*	intact	63 + 164
OR51B2	C > T	358	R > C	120	rs7952293	225	*DpnI*	intact	56 + 169
OR51F1	C > Del(1)	274	R > frameshift	92	rs34672924	209	*BsaAI*	intact	64 + 145
OR51G1	G > A	371	R > H	124	rs34742470	227	*DraIII*	pseudo	83 + 144
OR51J1	G > A	299	C > Y	100	rs1909261	188	*DpnI*	pseudo	26 + 162
OR51Q1	C > T	706	R > stop	236	rs2647574	173	*BlpI*	intact	80 + 93
OR52B4	C > Del(1)	119	A > frameshift	40	rs11310407	225	*BslI*	intact	80 + 145
OR52D1	A > T	662	Y > F	221	rs7950082	157	*NcoI*	intact	25 + 132
OR52H1	G > A	389	R > H	130	rs1566275	274	*BclI*	pseudo	79 + 195
OR52L1	C > T	418	R > C	140	rs4436525	176	*HhaI*	intact	25 + 151
OR52N4	A > T	514	R > stop	172	rs4910844	249	*PstI*	intact	90 + 159
OR52R1	C > T	386	T > I	129	rs7941731	186	*HpaII*	intact	23 + 163
OR6J1	G > A	362	R > H	121	rs3751484	182	*AclI*	intact	21 + 161
OR6Q1	C > Del(1)	685	L > frameshift	229	rs34846253	204	*TspRI*	intact	89 + 115
OR7C2	G > A	365	R > H	122	rs11883178	289	*DraIII*	pseudo	137 + 152
OR8B4	T > C	532	C > R	178	rs4057749	175	*AciI*	pseudo	20 + 155
OR8D2	G > A	365	R > H	122	rs2512219	188	*BclI*	pseudo	20 + 168
OR8G1	C > G	777	Y > stop	259	rs4268525	238	*RsaI*	intact	99 + 139
OR8J2	C > T	190	R > stop	64	rs55684887	221	*BslI*	intact	77 + 144
OR8K3	G > T	365	R > L	122	rs960193	237	*AciI*	intact	69 + 168
OR10A6	C > T	860	P > L	287	rs4758258	175	*AciI*	intact	44 + 131
OR10C1	C > T	163	Q > stop	55	rs17184009	202	*BsrI*	intact	62 + 140
OR10X1	G > A	198	W > stop	66	rs863362	201	*SmlI*	pseudo	51 + 150
OR12D2	G > T	360	R > L	120	rs2073153	298	*BsrBI*	intact	123 + 175

[a]Positions of polymorphisms in intact allele or pseudogene (pseudo) allele of the human OR genes (nt, nucleotide position; AA, amino acid position) are relative to the translation start site. Del(1) and Del(2) refer to deletions of 1 and 2 nucleotides, respectively.

[b]Identification number of the polymorphism in the single nucleotide polymorphism (SNP ID) database at the National Center for Biotechnology Information.

[c]The PCR product size (bp), amplified using the primers (see Table 2), the restriction enzyme used to digest amplified DNA, the OR allele cut by the restriction enzyme, and the size of the DNA fragments produced from digestion, is indicated[1].

TABLE 2

Primer Sets for Amplifying Human Olfactory Genes

Gene	Forward primer	Length (nt)[a]	tm[b]	Reverse primer	Length (nt)[a]	tm[b]
OR1A1	TTATAGCCTAAAAGACGCAGTGAT	24	58.6	GAGGAGATTCTCTTGTTGAAGAGTT	25	58.7
OR1B1	TTTTGTGACCACCGGCCACTTCTT	24	69.7	CAGCTGAAGGCAAACGTAGAATAG	24	61.6
OR1E3	AGTTCCTGCTCCTGGGCCTTCG	22	66.3	ACACAAATACATAGGCATGTGGAG	24	60.2
OR1F1	ATGTATTTCGTTTTCATGTTCGTG	24	60.2	AGAAGTGAGTGATGGCATTGTCT	23	60.2
OR1P1	CATATTCCGGCTTTCCTTGTGCTCG	25	66.9	GAAACAATGCGGATATAAGAGACC	24	60.2
OR1S1	AATATTCAAACCAAGAGTCAATCCA	25	60.1	AGGGTGTGTGTCAGAGCAATAATA	24	60
OR2A9	CCCATGTACTTCTTCCTCTCACAC	24	61.3	CAAAACTCAAAAAGAGAAAGGTCG	24	60.6
OR2AG1	AGTGGGTCTCCTGAACTGCTCT	22	61.7	CAAAGGAGATGGTGTTTTCTCTG	23	60.2
OR2J1	CACTTCATTCCTCCTTTACTTTCT	24	57.3	ACATGTCCTAAGCACTTTCTGAAG	24	59
OR2L8	CTTCATTGGTATTTCATGTTCCC	23	59.6	AGGGATCTTGGACGTAGATAAGTG	24	59.9
OR2S2	ATCTGCTTCACTACCTCCTCAGTC	24	60.3	GATCACGGAGTACCTAAGGGG	21	59.8
OR2T11	ATCTGTACCACTGTCCCAAAACTC	24	60.8	TTCATCAGGACTGGGTATCTCA	22	59.6
OR3A1	AGTCGTCTCCTGTCCCGCAAGC	22	69.3	CTCATGCGGGTGCTGTAGGTGAGG	24	71
OR4A8	GGTGATGGCCTATGATCGCTAC	22	62.9	TCCAGGTCCAATAAAGGGTATATG	24	60.3
OR4C16	TGAGTTCATTCTGCTTGGATTGACT	25	63.5	TGGAAGTAGAGAGGCAAGTATCAG	24	59.1
OR4E2	CGAGATCTTTCTGCTGATCATG	22	60	GTCTGCCCTAGTGAGTGAACAGTA	24	59.9
OR4K3	CTCCCTTTGGTGATTAAACTTGCT	24	61.9	CAAAGAAAAGGGTCACAACAGTAA	24	59.6
OR4X1	ACTTCCTGAGAAGCCACAACTT	22	59.4	CCAATAAATCTCCTCATGGCA	21	59.9
OR4X2	GAGGTGCAGAGGGTTTGCTTT	21	62.9	CAGCAGATCTGAGATGAGTTTGG	23	61.3
OR5AL1	ATGCTGCTTCATCACATTTGTAGT	24	60.1	CAGTGAATCCATACACATACG	21	54
OR5AR1	AAGAAAACAGCTCAATGGTGACT	23	59.3	TGTGAAGCTGAGTGTCTGTTGTAA	24	60
OR5D13	TTCTGGTTTTCTTGTTCGTCTACA	24	60.2	AACCAAGTTCTCCAACAGTTTAGG	24	60
OR5G3	CACTTTTCTGGGTTGTGCTG	20	59.3	GAAAATTGTATGAGTCATGGTGCT	24	59.4
OR5H6	ATCTCTCTCTGAATGCATGGTACA	24	60.2	TAGCTGAATGCATAGTTCATTGGT	24	60.1
OR5L1	CACCTCACAGCTATCACTGTCTTC	24	60.4	TTTCTCAGGCTGTAGATCACCG	22	61.7
OR5R1	CCTTTGTTGACCTTTGTTACTCCT	24	60	AATGCAGACTCTTCTTGACATCAG	24	60
OR51B2	TTTATTCACTCCCTTTCTGTTGTG	24	59.6	AGCACGTGTGATAACATGAGATTT	24	60
OR51F1	AACAGCGTGATCCTGTTTGTCAT	23	63.1	TCCATAAAAGTGAATCCATGAAGA	24	59.8
OR51G1	CCTGTTTCACTCAGCTCTTCTTCA	24	62.3	GGAGTGGCAGTATTGGAAGC	20	59.7
OR51J1	AATACTGAAATTAGTCTTGAAGCTG	25	55.5	TCAGTACGATAATAGCTCCTGTGC	24	59.8
OR51Q1	ATTATTATCGTGGATCCTCTGCTC	24	59.9	CGATGAGTCATAGATACACCAACC	24	59.8
OR52B4	TTGCTGGGCATCCCTGGCCTA	21	71.1	ACGGAACCAGAAGATAGCTAAGG	23	60.2
OR52D1	GATTCCATTCTCATTGCCATTTCCC	25	67.5	GGGTGAGGAAGGAGAAGAAGG	21	61.5
OR52H1	CCCAGGATGCCTTACACAAATGTTC	25	66.6	GGCAACACCTATATGCTCACAGTA	24	60.4
OR52L1	TACTTGTGGCCATGGCTCTGGAG	23	67.7	GGTGGCTTGGCAGAAGATAAGTTT	24	63.9
OR52N4	TAATTGCAAAGGTTGGGACTGC	22	63.4	GGAGAATCATGGTATAGGAGTTGG	24	60.1
OR52R1	TCCTCTCCTCCTCCACTCAACCTA	24	64.8	TAGAGTGTCGGAGTGGGAAGCCG	23	68.9
OR6J1	CTGACGGTCATGTCCTATGAAC	22	59.5	TTAATGATATTGGAGCCACAGAAG	24	59.5
OR6Q1	ACTGTGGATTTCCTGGTGTCTCT	23	61.3	CACCTTGGTCTGGACATACATAAA	24	60.2
OR7C2	TCCCAAAGATGCTGGTGAATATC	23	62.4	ATTTCCATATTTGTGCAGAAGGAC	24	60.6
OR8B4	CAACGTCATTGACCATTATCCG	22	62.4	AGGAATACAGAGGATGTTGGAGAG	24	60
OR8D2	TTCTGACAGCCATGGAATATGATC	24	62.4	ATGACTGACCGTATGAGACCTACA	24	60
OR8G1	GACCATCCTTTGCTCTTACATCTT	24	60	TATGTAGCTCAGAGGGTTCAACAT	24	59.2
OR8J2	CAGGGAACCTGGGCATCATC	20	65	GAAAATCTCAGCCACAATGAAAAC	24	61.2
OR8K3	ACACAGCTAGCTTTCTTTCTTGT	23	56.8	GAAATGACTAATGACGTTGTAGCC	24	59.1
OR10A6	TCACCTCACATCTGTGACGCTATT	24	62	CAAAGCCCTCTTCATCTCACTATT	24	60.2
OR10C1	GGTGACTGAGTTTCTTCTTCTCG	23	59.6	AGACGTATAGCCAATCTCCAAGG	23	60.8
OR10X1	TTGTCTCTACCTTCTCACCCTTGC	24	63.1	AAGCTACAACCTGTGACTGAAATG	24	59.8
OR12D2	GCTGCAGAACTTTCTCTCTACACAC	25	61	ATCACAGAGAAAATGATGGATACG	24	59.4

The PCR-amplified DNA can then be digested with the enzyme indicated in Table 1 to distinguish between pseudogene and intact alleles of each gene.

[a]The number of nucleotides that the primer contains.

[b]The melting temperature (tm) of the primer, which was calculated by the Primer3 program.

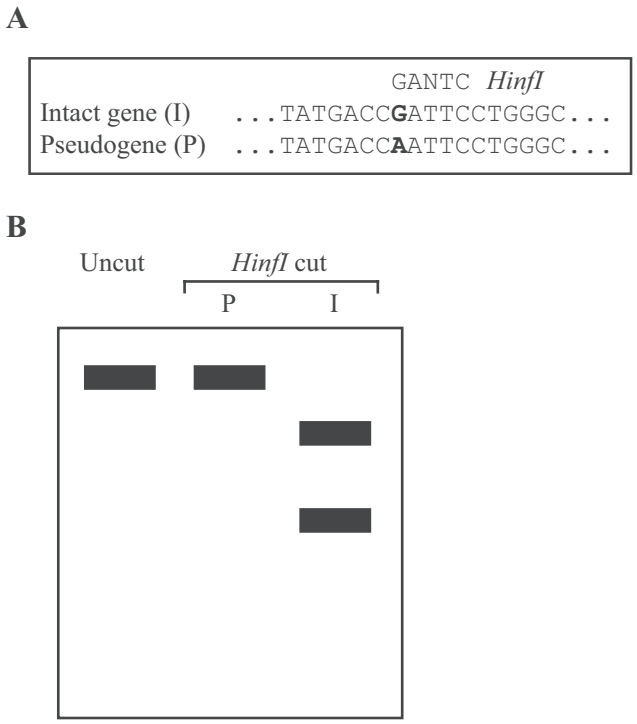

FIGURE 1

A single nucleotide polymorphism (G>A) in the OR3A1 gene can be used to distinguish between an intact gene and a pseudogene. (*A*) The intact gene contains a *HinfI* restriction site that is not present in the pseudogene. (*B*) DNA amplified from an intact gene or a pseudogene of OR3A1 can be distinguished by digestion with *HinfI* and electrophoresis on an agarose gel. Uncut DNA (Uncut) and DNA from the pseudogene allele will be the same size, but DNA from the intact allele will be digested into two smaller DNA fragments.

nucleotide 374, whereas the pseudogene allele has an "A", which results in changing a highly conserved arginine (R) aa to glutamine (Q) in the predicted OR3A1 protein. This polymorphism also results in the presence of a *HinfI* restriction site in the intact allele that is not present in the pseudogene allele (Fig. 1A). Therefore, a PCR-amplified DNA fragment from an intact allele will be cut by *HinfI*, whereas a PCR-amplified DNA fragment from a pseudogene allele will not be cut (Fig. 1B).

Fig. 2 shows the results of a genotyping assay for the human OR3A1 gene. For this assay, a 164-bp region of the OR3A1 gene was PCR-amplified from three individuals. A portion of the PCR-amplified DNA products was digested with the restriction endonuclease *HinfI*. Uncut (Lanes 1, 3, and 5) and cut (lanes 2, 4, and 6) products were then separated by electrophoresis using a 2% agarose gel. The PCR product from the intact allele was cut into DNA fragments of 119 bp and 45 bp (compare lane 6 with lane 5), whereas the PCR product from the pseudogene allele was uncut and remained 164 bp in length (compare lane 2 with lane 1). The 45-bp DNA fragment was undetectable, as it ran in the same location as what are likely primer dimers (asterisk). From this gel, it is also easy to determine the genotypes of the three individuals as P/P (lanes 1 and 2), I/P (lanes 3 and 4), and I/I (lanes 5 and 6) for the OR3A1 gene.

One of the major difficulties for further studies of OR function is that ligands for only 2–3% of mammalian ORs have been identified. The large number of segregating human OR pseudogenes is naturally occurring gene knockouts and therefore, provides a unique opportunity to perform genotype/phenotype studies to help identify ligands for these ORs. As the genotype assay outlined in this paper can be used to determine genotypes for most of the segregating human OR pseudogenes, >60 human OR genes could potentially be used for genotype/phenotype studies.

There are several advantages to this genotyping assay:

1. This assay uses well-established methodologies (PCR amplification, restriction enzyme digestion, and agarose gel electrophoresis) that require equipment commonly found in biochemistry and molecular biology laboratories. Therefore, this assay can be used by even small laboratories, as it does not require expensive or specialized equipment.

2. Although this assay was developed to distinguish between polymorphisms in the segregating human OR genes, it can be adapted easily for use with other genes or even to distinguish polymorphisms in noncoding DNA.

3. The source of DNA for large population-based studies has traditionally been whole blood, as the yield of

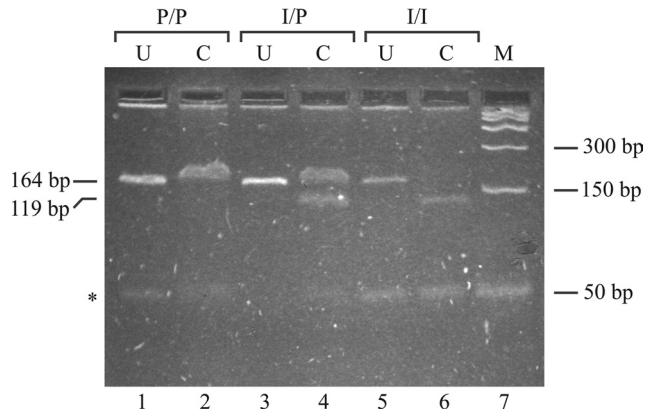

FIGURE 2

Genotypes of three individuals for the human OR3A1 gene. There are two alleles for the OR3A1 gene, a pseudogene allele and an intact allele. To distinguish among the three possible genotypes, P/P, I/P, and I/I, human cheek cell DNA was used in PCR reactions to amplify a 164-bp region of the OR3A1 gene. A portion of the amplified DNA was digested with *Hinf1*, and uncut (U; lanes 1,3, and 5) and cut (C; lanes 2, 4, and 6) DNAs were separated on a 2% agarose gel. The asterisk indicates primer dimers or a mixture of the 45-bp restriction fragment and primer dimers. Lane 7 contains DNA size markers (M).

DNA can be high. However, obtaining blood samples requires trained technicians and may be difficult to obtain when study subjects are not in proximity to one another. Many potential study subjects may also not want to participate, as they do not want to provide blood samples. In addition, blood samples must be refrigerated and processed within ~1 week.[20,21] Unlike procedures that isolate DNA from human blood, obtaining cheek cells by mouthwash is noninvasive. A large number of studies have shown that high-quality DNA with high yields can be obtained from cheek cells.[20–23] In addition, cheek cell mouthwash samples are stable at room temperature for at least 2 weeks.

In conclusion, the entire genotyping assay, from isolating an individual's DNA to determining the genotype, can be accomplished easily in a single day. DNA isolation requires only 20 min. For most of the segregating human OR pseudogenes, the PCR fragments amplified are small, between 150 bp and 250 bp, so amplification and electrophoresis times are short. Restriction endonucleases can be added directly to the PCR products, and DNA can be loaded directly onto agarose gels after digestion. It is therefore possible to process a large number of samples each day. This is important, as working with samples from a substantial number of subjects is required to detect some of the polymorphisms that are present at a low frequency in the human population. As the genotype assay presented in this paper is composed of a few straightforward steps, it is possible to automate them to produce a high-throughput screening system.

ACKNOWLEDGMENTS

This work was supported by grant 0516051091 from the Collegiate Science and Technology Entry Program of the New York State Department of Education and grant 1R25GM62003 from the Bridges to the Baccalaureate Program of the National Institute of General Medical Sciences. The authors thank Gary Sarinsky for his help with this project. We also thank Theodore Markus, Farshad Tamari, Loretta Taras, and Arthur Zeitlin for critical reading of the manuscript.

REFERENCES

1. Demaria S, Ngai J. The biology of smell. *Cell Biol* 2010;191: 443–452.
2. Buck L, Axel R. A novel multigene family may encode odorant receptors: a molecular basis for odor recognition. *Cell* 1991;68: 175–187.
3. Ngai J, Dowling MM, Buck L, Axel R, Chess A. The family of genes encoding odorant receptors in the channel catfish. *Cell* 1993;72:657–666.
4. Glusman G, Yanai I, Rubin I, Lancet D. The complete human olfactory subgenome. *Genome Res* 2001;11:685–702.
5. Zhang X, Firestein S. The olfactory receptor gene superfamily of the mouse. *Nat Neurosci* 2002;5:124–133.
6. Zhao H, Ivic L, Otaki JM, Hashimoto M, Mikoshiba K, Firestein S. Functional expression of a mammalian odorant receptor. *Science* 1998;279:237–242.
7. Araneda RC, Kini AD, Firestein S. The molecular receptive range of an odorant receptor. *Nat Neurosci* 2000;3:1248–1255.
8. Malnic B, Hirono J, Sato T, Buck LB. Combinatorial receptor codes for odors. *Cell* 1999;96:713–723.
9. Touhara K, Sengoku S, Inaki SK, et al. Functional identification and reconstitution of an odorant receptor in single olfactory neurons. *Proc Natl Acad Sci USA* 1999;96:4040–4045.
10. Saito H, Kubota M, Roberts RW, Chi Q, Matsunami H. RTP family members induce functional expression of mammalian odorant receptors. *Cell* 2004;119:679–691.
11. Keller A, Zhuang H, Chi Q, Vosshall LB, Matsunami H. Genetic variation in a human odorant receptor alters odour perception. *Nature* 2007;449:468–472.
12. Zozulya S, Echeverri F, Nguyen T. The human olfactory receptor repertoire. *Genome Biol* 2001;2:0018.1–0018.12.
13. Menashe I, Man O, Lancet D, Gilad Y. Different noses for different people. *Nat Genet* 2003;34:143–144.
14. Menashe I, Abaffy T, Hasin Y, et al. Genetic elucidation of human hyperosmia to isovaleric acid. *PLoS Biol* 2007;5:2462–2468.
15. Safran M, Chalifa-Caspi V, Shmueli O, et al. Human gene-centric databases at the Weizmann Institute of Science: GeneCards, UDB, CroW 21 and HORDE. *Nucleic Acids Res* 2003; 31:142–146.
16. Neff MM, Turk E, Kalishman M. Web-based primer design for single nucleotide polymorphism analysis. *Trends Genet* 2002;18: 613–615.
17. Vincze T, Posfai JRJ, Roberts RJ. NEBcutter: a program to cleave DNA with restriction enzymes. *Nucleic Acids Res* 2003; 31:3688–3691.
18. Rozen S, Skaletsky HJ. Primer3 on the WWW for general users and for biologist programmers. In Krawetz S, Misener S (eds): *Bioinformatics Methods and Protocols: Methods in Molecular Biology*. Clifton, NJ, USA: Humana, 2009;365–386.
19. Bloom MV. Human DNA fingerprinting by polymerase chain reaction. In Goldman CS (ed): *Tested Studies for Laboratory Teaching*, vol 15, *Proceedings of the 15thWorkshop/Conference of the Association for Biology Laboratory Education (ABLE)*, 1994; 1–13.
20. Quinque D, Kittler R, Kayser M, Stoneking M, Nasidze I. Evaluation of saliva as a source of human DNA for population and association studies. *Anal Biochem* 2006;353:272–277.
21. Hansen TO, Simonsen MK, Nielsen FC, Hundrup YA. Collection of blood, saliva, and buccal cell samples in a pilot study on the Danish nurse cohort: comparison of the response rate and quality of genomic DNA. *Cancer Epidemiol Biomarkers Prev* 2007;16:2072–2076.
22. Heath EM, Morken NW, Campbell KA, Tkach D, Boyd EA, Strom DA. Use of buccal cells collected in mouthwash as a source of DNA for clinical testing. *Arch Pathol Lab Med* 2001; 125:127–133.
23. Feigelson HS, Rodriguez C, Robertson AS, et al. Determinants of DNA yield and quality from buccal cell samples collected with mouthwash. *Cancer Epidemiol Biomarkers Prev* 2001;10:1005–1008.